

On comparing JPEG2000 and Intraframe AVC

Mourad Ouaret^a, Frederic Dufaux^a, Touradj Ebrahimi^a

^aEcole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

ABSTRACT

In this work, a performance evaluation of AVC Intra and JPEG2000 in terms of rate-distortion performance is conducted. A rich set of test sequences with different spatial resolutions is used in this evaluation. Furthermore, the comparison is made with both the Main and High profiles of AVC Intra.

For high spatial resolution sequences, our results show that JPEG2000 is very competitive with AVC High Profile Intra and outperforms the Main Profile. For Intermediate and low spatial resolution sequences JPEG2000 is outperformed by both Profiles of AVC Intra.

Keywords: Image Coding, AVC, H.264, MPEG-4, JPEG2000.

1. INTRODUCTION

AVC (H.264, MPEG-4 Part 10) [1], for Advanced Video Coding, is a digital video codec standard which is noted for achieving very high data compression. It has been developed by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG) as the product of a collective partnership effort known as the Joint Video Team (JVT). AVC [1] provides good video quality at substantially low bit rates. It is based on a block-based integer DCT transform. In addition, it performs spatial prediction for Intra frame coding and temporal motion estimation for inter frame coding to improve the compression efficiency.

On the other hand, JPEG2000 [2] is a wavelet-based compression standard for still images. It is created by the Joint Photographic Experts Group (JPEG) committee. Besides offering a number of new functionalities, it outperforms the original DCT-based JPEG standard in terms of compression efficiency in many situations.

Intra frame coding results in lower complexity encoders compared to Inter frame coding. In addition, it is suitable for random access, browsing and editing of video content since each frame is encoded on its own without using any information from its neighbouring frames. Furthermore, Intra frame coding does not allow error propagation which is attractive for error resilience. Due to these benefits Intra coding can be used for several applications such as digital cinema, satellite and medical imaging as well as video surveillance.

We will refer to AVC Main Profile as AVC MP and to AVC High Profile as AVC HP.

In [3], a performance evaluation of AVC MP Intra and JPEG2000 was conducted. It is reported that AVC Intra performs better than JPEG2000 in terms of rate-distortion performance for low and intermediate resolution sequences. The gain of AVC Intra over JPEG2000 in PSNR was around 0.5 ~ 2.0 dB. On the other hand, JPEG2000 performed better for higher resolution sequences with a gain around 0.5 ~ 1.0 dB in PSNR.

Furthermore, [4] compared AVC HP Intra and JPEG2000 for monochromatic still image coding. It is shown that their performance are identical. Nevertheless, JPEG2000 has a gain of 1 dB in PSNR over AVC HP Intra if the 8x8 transform is disabled for the encoder. However, the evaluation was performed on a small set of images, which reduces its consistency.

Finally, [5] performed the same comparison as [4]. However, [5] used video sequences at high resolutions instead of fixed resolution still images. The experimental results in [5] show that AVC HP Intra offers rate-distortion gain around 0.2 ~ 1.0 dB in PSNR over JPEG2000.

In this paper, the performance of JPEG2000 and AVC Intra (Main and High profiles) are evaluated for a richer set of sequences with various spatial resolutions. This paper is structured as follows. First, both standards are introduced and their compression algorithms are briefly described in Section 2. Then, the set of video sequences and the encoder parameters used are specified in Section 3. The rate-distortion curves are presented in Section 4. Finally, we draw some concluding remarks in Section 5.

2. COMPRESSION ALGORITHMS

2.1 AVC Intra

AVC [1] MP Intra is based on the block-based integer DCT transform. Unlike its predecessors, the block size for the transform is reduced from 8x8 to 4x4. AVC Intra takes advantage of the spatial correlation to improve the coding efficiency. The Intra coding of a macroblock consists in four main steps, spatial prediction, 4x4-DCT transform, scalar quantization, and entropy coding. If a macroblock is encoded in intra predictive mode, a prediction macroblock is formed based on previously encoded and decoded macroblocks. This prediction macroblock is subtracted from the current macroblock prior to encoding. For the luminance samples, the prediction macroblock is computed for each 4x4 subblock or for a 16x16 macroblock. There are a total of 9 optional prediction modes for each 4x4 luminance subblock, 4 optional modes for a 16x16 luminance macroblock and four modes for each 8x8 chrominance macroblock.

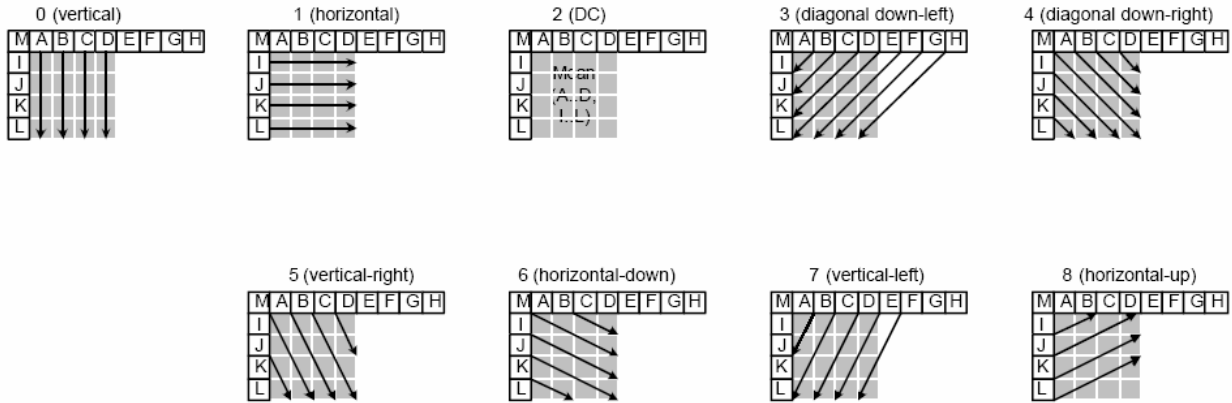


Fig.1. 4x4 luminance prediction modes.

The arrows in Figure 1 indicate the direction of prediction in each mode. For modes 3 to 8, the predicted samples are formed from a weighted average of the prediction samples A-M. For mode 2, the predicted samples are formed from the mean of samples A to D and I to L. For modes 0 and 1, the samples are predicted vertically and horizontally from samples A to D and I to L respectively. The encoder selects the prediction mode for each block that minimizes the prediction residual.

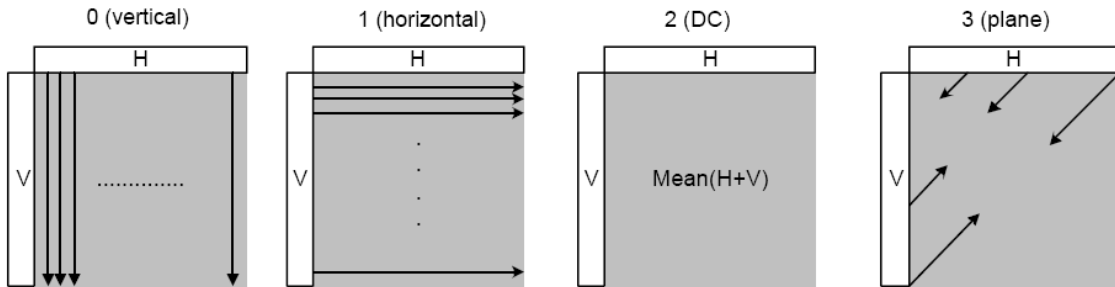


Fig.2. 16x16 luminance prediction modes.

Figure 2 shows the four modes for the 16x16 luminance macroblocks. In mode 3, a linear plane function is fitted to the upper and left-hand samples H and V. The four chroma prediction modes are very similar to the 16x16 luma prediction modes, except that the order of mode numbers is different: DC (mode 0), horizontal (mode 1), vertical (mode 2) and plane (mode 3).

Then, the macroblocks are transformed using a 4x4 separable integer transform with properties similar to that of DCT. Furthermore, a non-uniform scalar quantization is used to quantize the transformed coefficients where each macroblock has a Quantization Parameter (QP). Finally, the quantized coefficients are entropy coded. AVC Intra supports two modes

for entropy coding: Context-Adaptive Variable Length Coding (CAVLC) and Context-Adaptive Binary Arithmetic Coding (CABAC). With CAVLC, multiple VLC tables are available and the encoder switches among them based on previously encoded syntax elements. On the other hand, CABAC is based on an arithmetic coder. Using arithmetic coding, each symbol of the alphabet can be assigned a non-integer number of bits, therefore outperforming VLC tables. Furthermore, a context model is build based on the statistics of previously transmitted syntax elements in order to better estimate conditional probabilities. This allows for the adaptation to non-stationary statistics. For these reasons, CABAC achieves substantially better coding performance when compared to CAVLC. However, it requires significantly higher computational complexity.

Further, the AVC MP is extended by introducing the High Profile encoder or AVC Fidelity Range Extensions (FR Ext). In the latter, an 8x8 integer transform is introduced. The encoder chooses adaptively between the 4x4 and the 8x8 transform for the luminance samples. More precisely, the 8x8 transform is used if and only if the 8x8 spatial prediction is used. Using the 8x8 transform enables the preservation of fine details and textures which generally require larger basis functions. Furthermore, three sets of context models are added in CABAC for the 8x8 transform coding. Meanwhile, CAVLC is used by regrouping the 8x8 transform coefficients into groups of 4x4. Furthermore, AVC HP allows using more than 8 bits per sample for more accuracy. In addition, The High Profile supports higher color space resolutions such as YUV 4:2:2 and YUV 4:4:4 with interesting features such as scaling matrices for perceptually tuned and frequency-dependent quantization specified at the encoder, the reversible residual color from RGB to YC_gC_o transform which is applied only to residual data and the lossless coding capability. For more details on AVC please refer to [1], [6].

2.2 JPEG2000

The JPEG2000 [2] standard makes use of the Discrete Wavelet Transform (DWT). JPEG2000 supports some important features such as improved compression efficiency, lossless and lossy compression, multi-resolution representation, Region Of Interest (ROI) coding, error resilience and a flexible file format. Figure 3 depicts the JPEG2000 fundamental building blocks.



Fig.3. the JPEG2000 fundamental building blocks.

In the Pre-Processing stage, an inter-component transformation is used to decorrelate the color data. There are two possible transforms. Both transforms operate on the first three components of an image tile with the implicit assumption that these components correspond to red–green–blue (RGB). One transform is the Irreversible Color Transform (ICT), which is identical to the traditional RGB to YC_bC_r color transformation and can only be used for lossy coding. The other transform is the Reversible Color Transform (RCT), which is a reversible integer-to integer transform that approximates the ICT for color decorrelation and can be used for both lossless and lossy coding. Then, the DWT is applied to the processed samples. The DWT provides a multi-resolution image representation. Furthermore, it achieves better compression due to its good energy compaction and ability to decorrelate the image across a larger scale. The resulting wavelet coefficients are quantized using a uniform quantizer with a central deadzone. It is shown that this quantizer is optimal for a continuous signal with a Laplacian distribution such as DCT or wavelet coefficients. The coefficients are gathered in subbands. Each subband is partitioned into small rectangular blocks called codeblocks where each codeblock is independently coded by the Adaptive Binary Arithmetic encoder. Finally, the output of the arithmetic encoder is organized as a compressed bit-stream which offers a significant degree of flexibility. This enables features such as random access, region of interest coding, and scalability. This flexibility is achieved partly through the various structures of components, tiles, subbands, resolution levels, and codeblocks. For more details on the JPEG2000 standard refer to [2].

3. EVALUATION METHODOLOGY AND MATERIAL

3.1 Video sequences

Table 1 represents the set of video sequences used. The set used is rich as it contains a total of 26 sequences representing 11 main sequences at different spatial resolutions. The set contains sequences with high texture such as City, Mobile and Bus. On the other hand Harbour, Football, Soccer and Coastguard contain more or less uniform texture with significant motion. Furthermore, Crew contains uniform texture with sudden illumination changes due to camera flashes. The Hall monitor is a video surveillance sequence. Finally, Akiyo and Foreman are typical video conferencing sequences.

Sequence name	Frame rate	Resolution (4:2:0)
Soccer	15	176x144
	15	352x288
	30	704x576
City	15	176x144
	15	352x288
	30	704x576
Harbour	15	176x144
	15	352x288
	30	704x576
Mobile	7.5	176x144
	15	352x288
Hall monitor	15	176x144
	30	352x288
Foreman	7.5	176x144
	15	352x288
Football	7.5	176x144
	15	352x288
Coast guard	15	176x144
	30	352x288
Crew	15	176x144
	30	352x288
	30	704x576
Bus	7.5	176x144
	15	352x288
Akiyo	30	176x144
	30	352x288

Table 1. Video sequences used in the performance evaluation.

3.2 Encoder parameters

For the AVC Intra coding, the publicly available reference software (JM 10.2) [7] was used. Each frame of the sequence was coded as an I-frame with the following settings:

- Main and High Profile encoding.
- CABAC for Main and High Profile encoder.
- The 8x8 transform for High Profile encoding only.

The software KAKADU version 4.4 [8] was used for the JPEG2000 compression with the following settings:

- Codeblock size of 64x64.
- One tile per frame.
- 3, 4 and 5 decomposition levels for the resolutions 176x144, 352x288 and 704x576 respectively.
- Visual Frequency Weighting switched-off.

3.3 Evaluation

Rate-distortion curves of the luminance component for the sequences are computed in the following way, The AVC encoder was run with a fixed Quantization Step (QP). The AVC decoder then provided the bit rate and the PSNR data. Hence, the target bit rate per sample for the JPEG2000 encoder was computed using the bit rate from the AVC decoder. This way the rate-distortion points for both encoders were computed for similar bit rates.

4. PERFORMANCE EVALUATION

4.1 Rate-distortion plots

In this section, the rate-distortion plots for AVC Main Profile, High Profile and JPEG2000, for our set of video sequences are analyzed. The plots are clustered with respect to spatial resolutions.

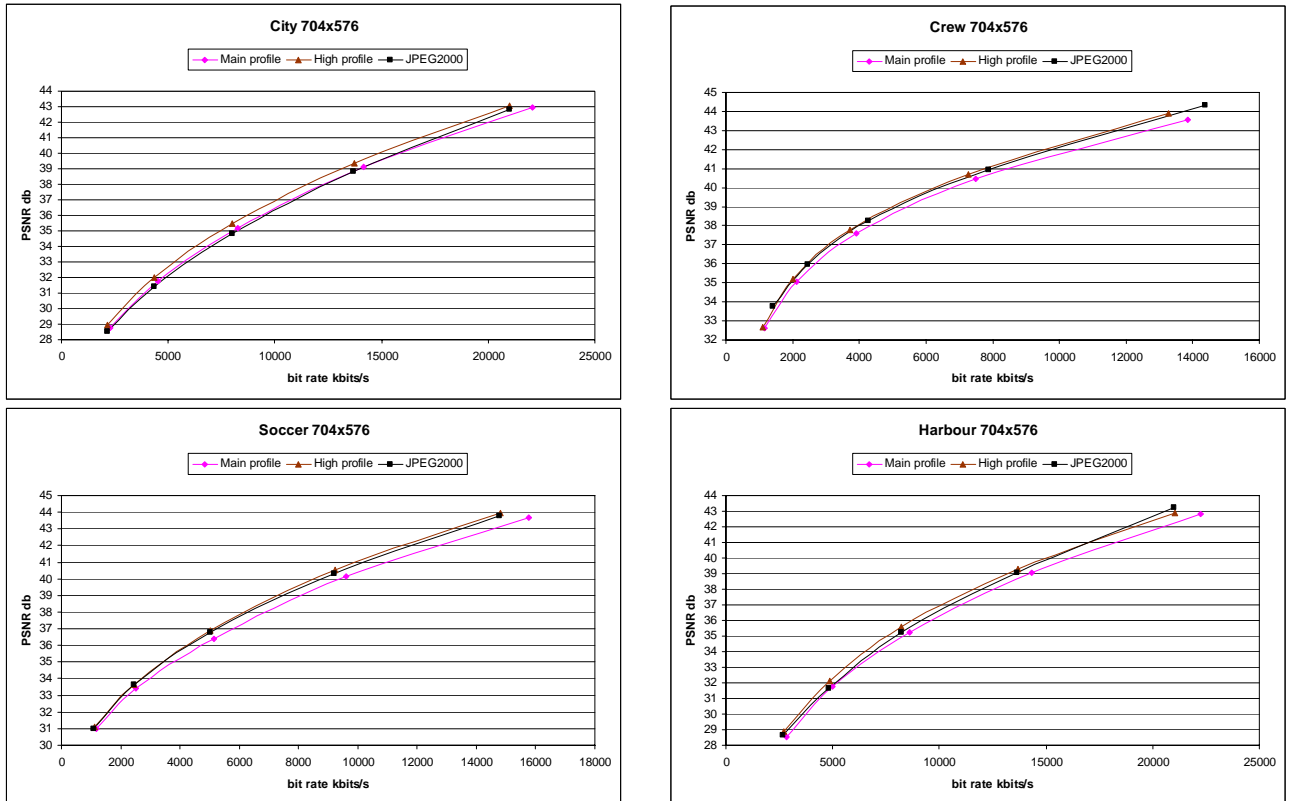


Fig.4. R-D plots for video sequences with spatial resolution 704x678.

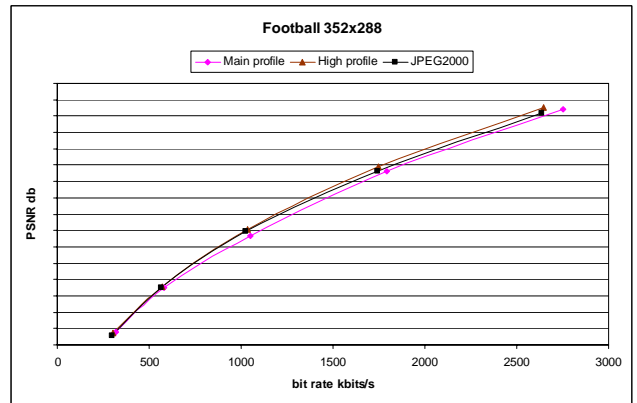
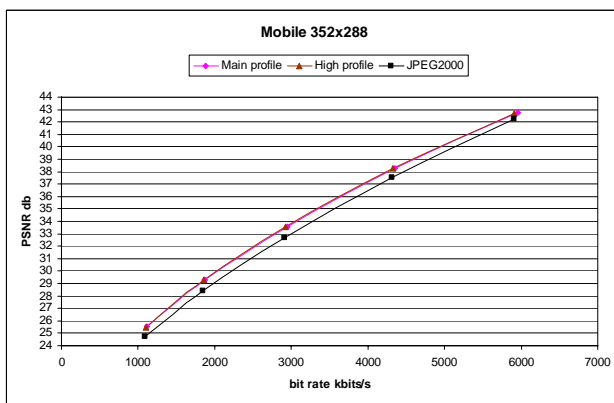
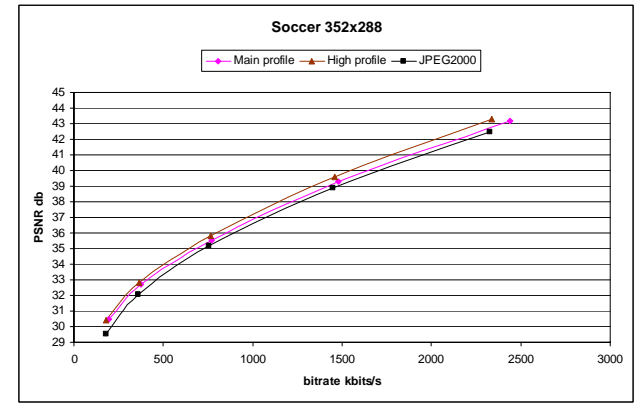
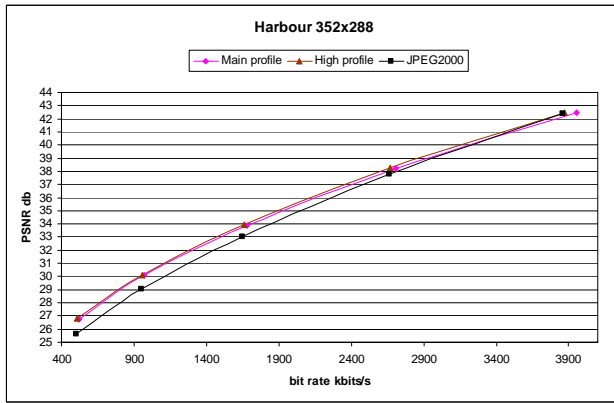
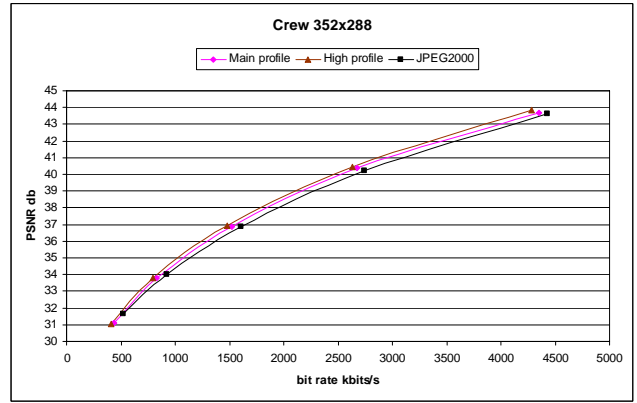
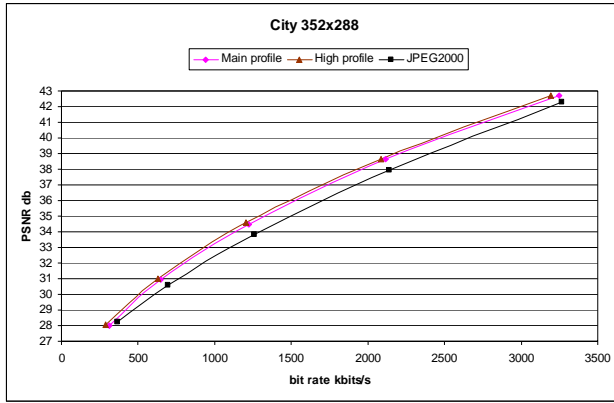


Fig.5. R-D plots for video sequences with spatial resolution 352x288.

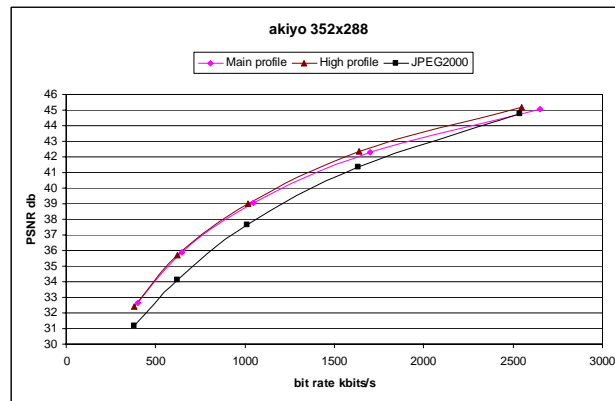
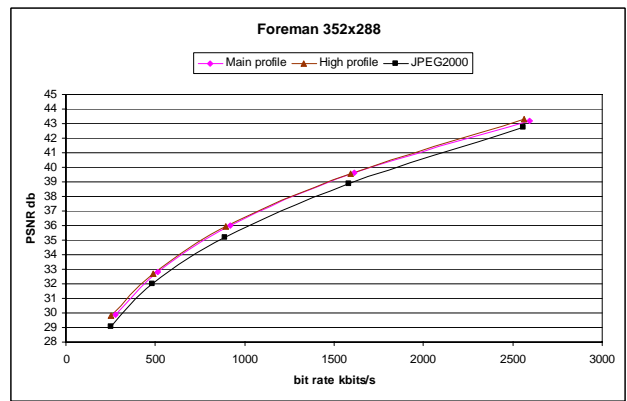
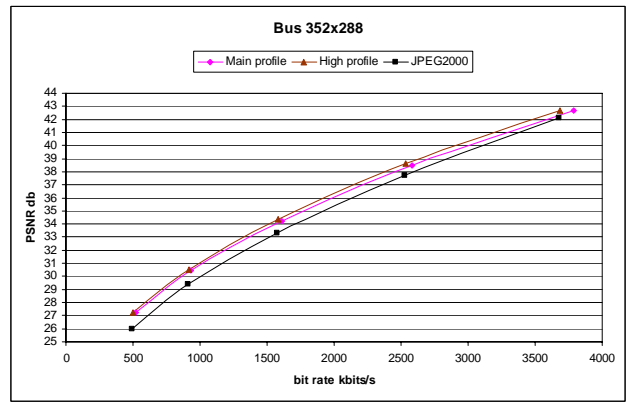
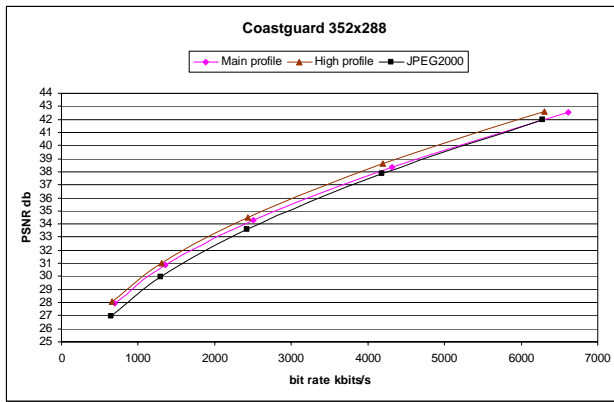
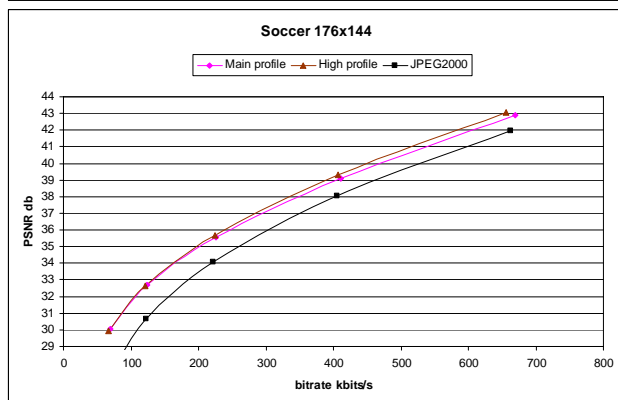
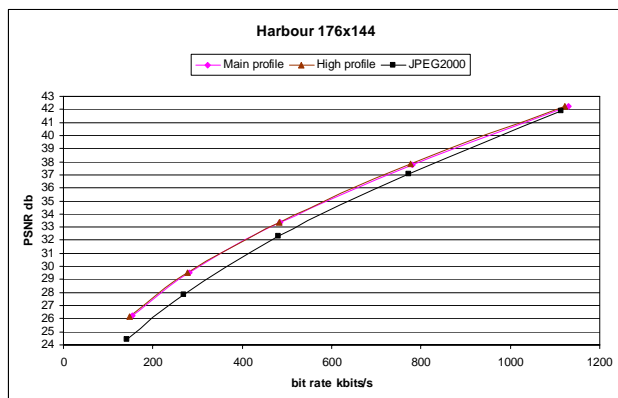
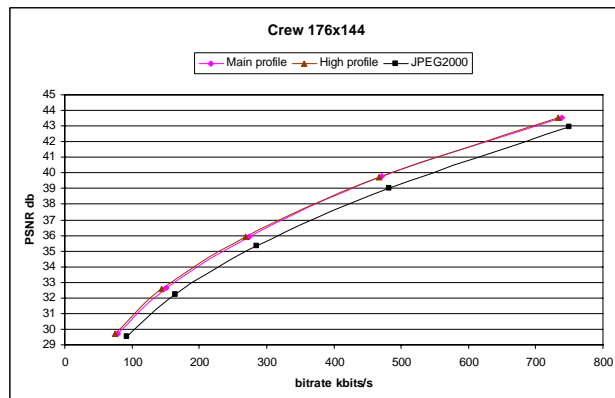
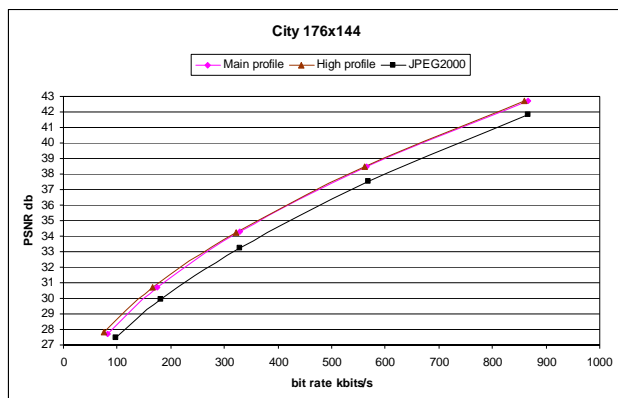


Fig.6. R-D plots for video sequences with spatial resolution 352x288.



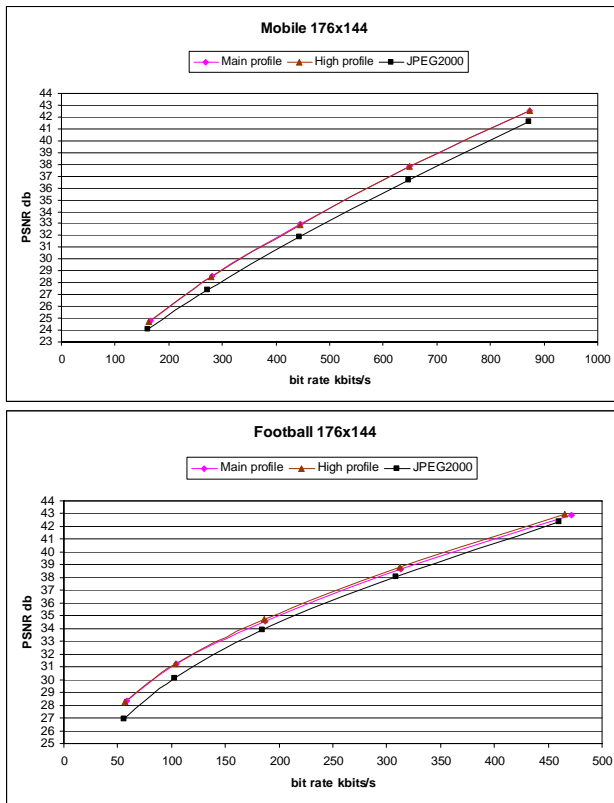
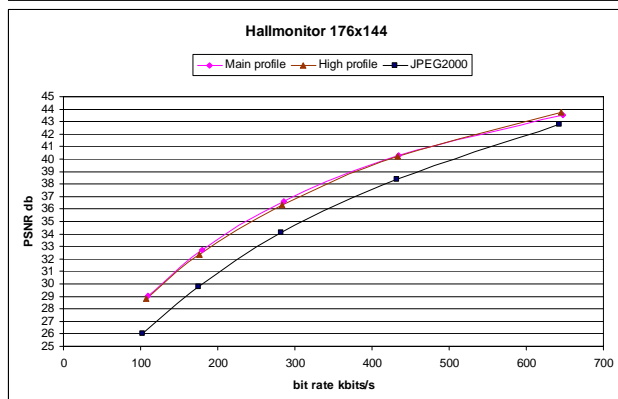
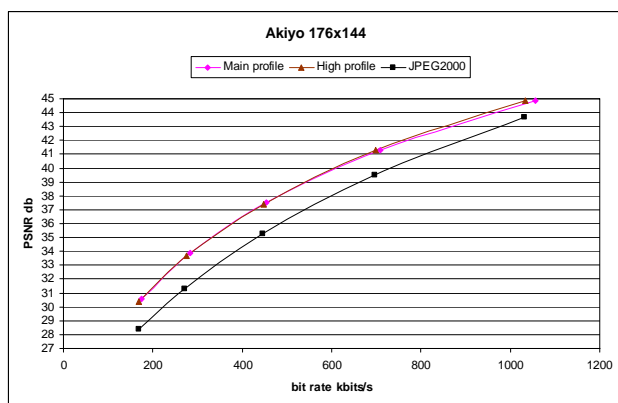
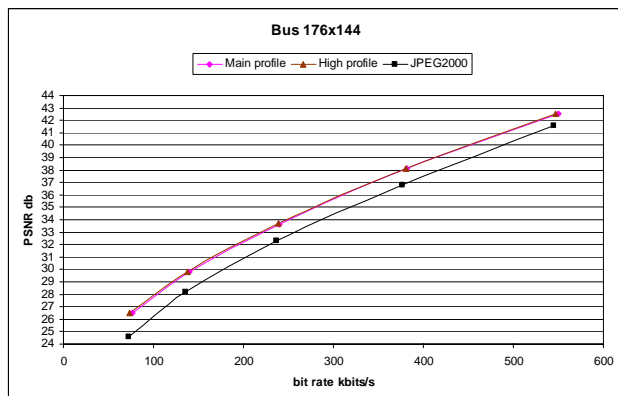
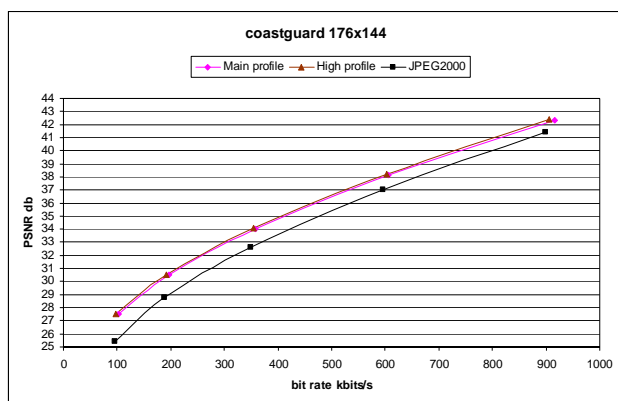


Fig.7. R-D plots for video sequences with spatial resolution 176x144.



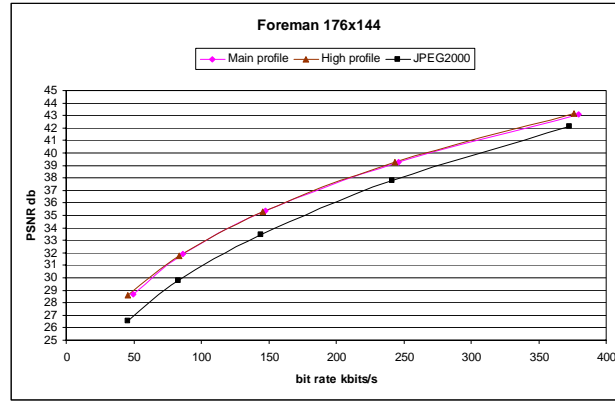


Fig.8. R-D plots for video sequences with spatial resolution 176x144.

For sequences with resolution equal to 704x576, it is reported in [5] that for the City sequence AVC HP Intra outperforms JPEG2000 with a gain around 1.0 dB in PSNR. This is not our case since the gain of AVC HP over JPEG2000 is around 0.6 dB at average bit rates and 0.3 dB at high bit rates. In addition, it is reported in [5] and [3] that AVC HP Intra outperforms JPEG2000 for the Harbour sequence with a gain around 0.8 ~ 1.0 db. In our case, the gain is around 0.2 in favor of AVC HP Intra at average bit rates and in favor of JPEG2000 at high bit rates. Finally, it is reported that AVC HP Intra has a gain of around 0.9 dB over JPEG2000 for the Crew sequence. In our case, the gain is around 0.1 dB in favor of AVC HP Intra. For the three sequences mentioned previously, JPEG2000 outperforms AVC MP Intra for the Harbour and Crew sequences with gains around 0.1 ~ 1.0 and 0.5 dB in PSNR respectively. Further, it performs as well as AVC MP Intra for the City sequence. Thus, JPEG2000 globally outperforms AVC MP Intra and is very close to AVC HP Intra in rate-distortion performance. The Soccer sequence confirms this as the performance of JPEG2000 is very close to AVC HP Intra, around 0.1 dB difference, and outperforms AVC MP Intra with around 0.6 gain in PSNR. This difference in rate-distortion performance compared to the same sequences in [5] and the Harbour sequence in [3] is due to the fact that [5] and [3] use the Visual Frequency Weighting for the JPEG2000 encoder, which results in a drop in rate distortion performance of JPEG2000. In [4], JPEG2000 is very close to AVC HP Intra for set of monochromatic still images with much higher resolution up to 2048x3072.

For sequences with resolution equal to 352x288, AVC Intra in both profiles outperforms JPEG2000. The gain is around 0.5 ~ 1.0 dB for the High Profile and around 0.1 ~ 0.5 for the Main Profile. Nevertheless, the difference between AVC Intra and JPEG2000 is small for sequences with more or less uniform texture such as Crew, Harbour, Crew, Football and Coastguard.

For sequences with resolution equal to 176x144, globally the Main and High Profiles have more or less the same rate-distortion performance since the High Profile has been introduced to improve the performance for high spatial resolutions. Furthermore, AVC Intra has a gain of 1.0 ~ 2.0 dB in PSNR over JPEG2000.

Finally, compared to [3] and [4], in addition to Visual Frequency Weighting mentioned previously, we use different wavelet decomposition levels for the different spatial resolutions. On the other hand, [3], [4] and [5] use a fixed number of 5 decomposition levels, which is not always optimal.

5. CONCLUSIONS

For high spatial resolution sequences, our results show that JPEG2000 is very competitive with AVC High Profile Intra with around 0.1 dB difference in PSNR in favor of AVC High Profile. Furthermore, JPEG2000 outperforms the Main Profile with gains around 0.1~1.0 dB in PSNR. For Intermediate and low spatial resolution sequences JPEG2000 is outperformed by both profiles of AVC Intra. Thus, JPEG2000 is interesting for applications with high resolution video such as satellite and medical imaging and digital cinema. In addition, JPEG2000 provides some interesting features such

as scalability, Region of Interest definition and rate control that AVC Intra does not provide. This is in addition to royalty fee free, license fee free nature of the JPEG2000 standard.

The evaluation methodology used to compare compression schemes is of primary importance. In this paper, we considered a simple evaluation methodology which consisted in comparison of rate-distortion characteristics of each codec under study for a number of video sequences of different resolutions. The distortion metric was chosen to be the widely used PSNR measure. The rate was obtained by first identifying the bitrate obtained with AVC intra encoder for a given set of appropriate parameters. Then, the JPEG 2000 encoder was run in order to produce the same amount of bits for the same video sequences. This approach heavily relies on the rate control mechanism of JPEG2000 encoder which is usually out of the scope of standardization, and could have an adverse effect in the performance of JPEG2000, as it takes into account the particular rate control mechanism employed in the codec used in our experiments. For instance, the amounts of bits spent per frame in JPEG2000 tests is imposed, which is an additional constraint when compared to AVC tests where the overall bit rate is imposed and the amount of bits per frame can vary.

A more appropriate evaluation methodology would be to create rate-distortion points at random points for all codecs under study within a given range and to either bypass or employ in both cases rate-control mechanisms. Other improvements consist in comparing the visual quality of decoded data as produced by different codecs. Furthermore, a comparison of complexity, memory requirements and power consumption between the codecs under study in this paper should be performed in order to produce a better understanding of their relative performance.

6. ACKNOWLEDGMENTS

This work was developed within DISCOVER, a European Project (www.discoverdvc.org), funded under the European Commission IST FP6 programme.

REFERENCES

1. Thomas Wiegand, Gary J. Sullivan, Gisle Bjøntegaard, and Ajay Luthra, *Overview of the H.264/AVC Video Coding Standard*, IEEE Trans. on Circuits and Systems for Video Technology, vol. 13, no. 7, (July 2003).
2. Skodras, A., Christopoulos, C., Ebrahimi, T., *The JPEG 2000 still image compression standard*, Signal Processing Magazine, IEEE vol. 18, Issue 5, (September. 2001) 36 - 58.
3. Detlev Marpe, Valeri George, and Hans L. Cycon, *Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in intra coding mode*, Wavelet Applications in Industrial Processing, Proceedings of the SPIE, Rhode Island, USA, October 2003.
4. Detlev Marpe, Stephen Gordon, and Thomas Wiegand, *H.264/MPEG4-AVC Fidelity Range Extensions: Tools, Profiles, Performance, and Application Areas*, IEEE International Conference on Image Processing, Genova, Italy, September 2005.
5. Pankaj Topiwala, *Comparative study of JPEG2000 and H.264/AVC FRExt I- frame coding on high-definition video sequences*, Optical Information Systems III, Proceedings of the SPIE, vol. 5909, (2005) 284-292.
6. Joint Video Team of ITU-T and ISO/IEC, *Draft Text of H.264/AVC Fidelity Range Extensions Amendment*, Doc. JVT-L047, September 2004.
7. <http://iphome.hhi.de/suehring/tml/>
8. <http://www.kakadusoftware.com/>